PB87-212742

International Sulfur Deposition Model Evaluation (ISDME)

(U.S.) Environmental Protection Agency
Research Triangle Park, NC

Jul 87

U.S. Department of Commerce
National Technical Information Service
NTIS

EPA/600/D-87/232
July 1987

THE INTERNATIONAL SULFUR DEPOSITION MODEL EVALUATION (ISDME)

By

Terry L. Clark,[1,4] Robin L. Dennis,[1] Steve Seilkop,[2] Mayer Alvo,[3]
Eva C. Voldner,[4] and Marvin P. Olson[4]

CR811563

EPA Project Officer
Carl Hayes

[1]Atmospheric Sciences Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711

[2]Analytical Sciences Incorporated, Research Triangle Park, NC 27711

[3]University of Ottawa, Ottawa, Ontario, Canada

[4]Atmospheric Environment Service, Downsview, Ontario, Canada

ATMOSPHERIC SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before comp'

PB87-212742

1. REPORT NO. EPA/600/D-87/232	2.
4. TITLE AND SUBTITLE THE INTERNATIONAL SULFUR DEPOSITION MODEL EVALUATION (ISDME)	5. REPORT DATE July 1987
	6. PERFORMING ORGANIZATION CODE --
7. AUTHOR(S) T.L. Clark and R.L. Dennis,[1] S. Seilkop,[2] M. Alvo,[3] Eva C. Voldner and M.P. Olson[4]	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS [1]U.S. EPA RTP, NC 27711 [2]Analytical Sciences Inc. RTP, NC 27711 3 University of Ottawa Ottawa, Ontario, Canada 4 Atmospheric Environment Downsview, Ontario, Can.	10. PROGRAM ELEMENT NO. N104/C/05/05-3166 (FY-87)
	11. CONTRACT/GRANT NO. CR811563
12. SPONSORING AGENCY NAME AND ADDRESS Atmospheric Sciences Research Laboratory-RTP, NC Office of Research and Development U.S. Environmental Protection Agency Research Triangle Park, North Carolina 27711	13. TYPE OF REPORT AND PERIOD COVERED
	14. SPONSORING AGENCY CODE EPA/600/09
15. SUPPLEMENTARY NOTES	

16. ABSTRACT

Eleven linear-chemistry atmospheric models of sulfur deposition were evaluated for each season of 1980. The evaluation data set consisted of sulfur wet deposition amounts calculated from screened precipitation chemistry measurements at 46 sites across eastern North America. The focus of this study differed substantially from those of preceding model evaluation studies in that the ISDME emphasized the ability of the models to replicate, within the uncertainties of the observations, the spatial patterns of observed seasonal amounts. Patterns of the predictions and observations were constructed via an interpolation technique known as simple kriging, which minimizes interpolation errors and estimates uncertainties resulting from the interpolation errors as well as measurement errors.

The evaluation results indicated that for all seasons but spring, the models generally did not mimic the observed location of the seasonal maximum amounts of sulfur wet deposition within the uncertainty limits. However, the interpolated predictions of eight models were within the uncertainty limits of the interpolated observations across at least 80% of the evaluation region for at least three seasons.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS (This Report) UNCLASSIFIED	21. NO. OF PAGES 24
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA 22161

Notice

The information in this report has been funded in part by the United States Environmental Protection Agency. The report has been subject to the Agency's peer and administrative review and has been approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

To be presented at the 80th Annual Meeting of the Air Pollution Contro Association, New York, NY, June 21-26, 1987.

THE INTERNATIONAL SULFUR DEPOSITION MODEL EVALUATION (ISDME)

Terry L. Clark* and Robin L. Dennis*
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina

Eva C. Voldner and Marvin P. Olson
Atmospheric Environment Service
Downsview, Ontario, Canada

Steve Seilkop
Analytical Sciences Incorporated
Research Triangle Park, North Carolina

Mayer Alvo**
University of Ottawa
Ottawa, Ontario, Canada

ABSTRACT

Eleven linear-chemistry atmospheric models of sulfur deposition were evaluated for each season of 1980. The evaluation data set consisted of sulfur wet deposition amounts calculated from screened precipitation chemistry measurements at 46 sites across eastern North America. The focus of this study differed substantially from those of preceding model evaluation studies in that the ISDME emphasized the ability of the models to replicate, within the uncertainties of the observations, the spatial patterns of observed seasonal amounts. Patterns of the predictions and observations were constructed via an interpolation technique known as simple kriging, which minimizes interpolation errors and estimates uncertainties resulting from the interpolation errors as well as measurement errors.

The evaluation results indicated that for all seasons but spring, the models generally did not mimic the observed location of the seasonal maximum amounts of sulfur wet deposition within the uncertainty limits. However, the interpolated predictions of eight models were within the uncertainty limits of the interpolated observations across at least 80% of the evaluation region for at least three seasons.

---

* Also affiliated with the U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

** Also affiliated with the Systems and Informatics Directorate, Environment Canada.

INTRODUCTION

In Europe and North America regional, linear-chemistry models have been developed to estimate long-term air concentrations and dry/wet deposition amounts of sulfur oxides. The results of these models, which do not always concur, have been used to determine spatial patterns and source-receptor relationships, assess emission control strategies and estimate sulfur budgets. Because of their wide-spread use and the significance of decisions based on the results, it is necessary to i) rigorously evaluate and ii) compare the performance of these models to better understand their behavior.

Previous model evaluations were not rigorous for several reasons. First, as was the case in North America, the paucity of precipitation chemistry sites precluded statistically meaningful evaluations. Second, previous evaluations generally relied on only distributional statistics (e.g., means, biases, root-mean-square errors, etc.), which compared attributes of ensemble predictions and observations at points. These model performance measures are severely limited, since they do not consider spatial characteristics of the data. That is, these measures cannot quantify the degree to which the models replicate spatial patterns of the observations. Moreover, when distributional statistics are used exclusively, those models which predict flat fields for parameters which exhibit high spatial variability could deceivingly appear to perform well.

Furthermore, with few exceptions, it was impossible to compare the performance of the various models, since each independent evaluation used different sets of model input and evaluation data, as well as different performance measures. In addition, each evaluation study used independently derived methods to calculate amounts of sulfur wet deposition.

Therefore, as recommended by the Atmospheric Science and Analysis Work Group, established under the U.S./Canadian Memorandum of Intent on Transboundary Air Pollution (MOI, 1982), the U.S. Environmental Protection Agency and Environment Canada cosponsored the International Sulfur Deposition Model Evaluation (ISDME) to more rigorously evaluate, as well as intercompare, the performance of these linear-chemistry models. This paper summarizes the evaluation approach and results and compares the performances of eleven such models. The study is discussed in detail in the final report (ISDME, 1987).

ISDME APPROACH

The goals of this model evaluation were established on the basis of i) the variables predicted by the models; ii) how the user community applies the results of these models and iii) what data, as well as the spatial and temporal scales of the data, were available to evaluate the models. Simply stated, the goals were:

1) For each season of one year, assess each model's ability to replicate the observed spatial and temporal patterns of sulfur wet deposition within the uncertainties of the evaluation data.

2) Delineate for each model, areas of significant differences between the observed and predicted amounts of sulfur wet deposition.

3) Assess the reasonableness of the seasonal predictions of air concentrations and dry deposition amounts.

4) Attempt to explain the differences in the model performances.

There were several unique aspects of the ISDME that represent improvements over the less rigorous model evaluation studies of the past. First, the ISDME did not rely solely on distributional statistics computed from pairs of observations and predictions at specific sites to assess model performance. Instead, the evaluations focused on the replication of spatial patterns as determined by an interpolation technique known as kriging (Finkelstein, 1984). This technique minimizes interpolation errors and quantifies the interpolation and measurement errors, which provide uncertainty limits of the evaluation data. Second, these uncertainty limits are used to determine whether the differences between predictions and observations are significant. Finally, instead of basing model performance solely on annual comparisons, overall model performance in the ISDME was based strictly on seasonal comparisons, which, by virtue of the finer temporal scale, afforded a more rigorous evaluation of the models. The seasonal evaluations also identified those seasons for which the models performed best and worst. Because preceeding studies focused on annual evaluations, these models were also evaluated for the annual period.

Modelers in Europe and North America were asked to apply their models in a "blind" testing mode using standardized sets of 1980 meteorological and sulfur dioxide emissions data for eastern North America. Predictions of seasonal and annual mean air concentrations and dry and wet deposition amounts at as many as 65 sites were submitted by three Canadian and six United States organizations. Of the eleven models (some organizations submitted predictions from more than one model), four were statistical, six were Lagrangian and one was a Lagrangian/Eulerian hybrid (Table 1). The attributes of these models are summarized in the final report (ISDME, 1987).

The year 1980 was selected as the evaluation year, since it was the only year when both a detailed emissions inventory and an adequate data base of precipitation chemistry data were available across eastern North America. However, an adequate quantity of air quality measurements and dry deposition estimates were not available to provide a comprehensive data base, hence the evaluation focused on sulfur wet deposition.

Seasonal and annual amounts of sulfur wet deposition were derived from 1980 precipitation chemistry data sampled at sites of the five major networks operating in North America during this year and archived in the Acid Deposition System (Watson and Olsen, 1984). These networks collected and analyzed data for daily, weekly or monthly periods. Amounts of sulfur wet deposition were calculated using a method recommended by the Unified Deposition Data Base Committee (UDDBC, 1985), a group comprised of Canadian and United States scientists.

The computed wet deposition amounts were classified into three levels of uncertainty based on site representativeness of the regional scale, data capture, precipitation collection efficiency, and contamination from sea salt (Voldner et al., 1984). Only data from the highest two levels comprised the evaluation data set. The number of sites passing this screening procedure ranged from 32 for the annual period to 46 for spring. Supplementing the wet deposition data, but only for comparison purposes, were the seasonal and annual means of sulfur dioxide and sulfate air concentrations from four regionally representative Canadian sites.

Using these evaluation data, model performances were assessed on the basis of the following elements:

1) Replication of seasonal spatial patterns of sulfur wet deposition amounts (i.e., pattern magnitude, position and shape) within data uncertainty limits;

2) Concurrence with the magnitude and location of the maximum observed seasonal amounts of sulfur wet deposition within the uncertainty limits;

3) Replication of mean temporal patterns (i.e., the distribution of the annual amounts of sulfur wet deposition across the seasons); and

4) Reasonableness of the predictions of seasonal mean air concentrations and dry deposition amounts of sulfur dioxide and sulfate.

Since the conventional model performance measures recommended by the American Meteorological Society (Fox, 1981) focused on distributional statistics, they alone were inappropriate for the stated evaluation elements. Consequently, the ISDME required different, more rigorous performance measures with an emphasis on spatial patterns.

MODEL PERFORMANCE MEASURES

Six ISDME model performance measures (see Table 2) were selected on the basis of their relevance to the evaluation elements listed above and their interpretability. The primary measure of pattern replication (Element 1) pertained to the area of significant differences between interpolated predictions and observations. Differences were deemed significant when the interpolated predictions fell outside the uncertainty range, as defined by $\pm$ two standard deviations about the interpolated observations. The values for this measure were equal to the percentage of half-degree cells within the model evaluation region and each of four sub-regions, depicted in Figure 1, where significant differences occurred. This measure was also useful in objectively identifying regions of good and poor agreement between the interpolated predictions and observations.

Two additional measures quantified differences between the observed and predicted spatial patterns -- the discrepancy between the centroids and the discrepancy between the orientation of the major axes. The centroids, which were calculated from the interpolated fields, are analagous to the center of mass, which is an attribute of pattern position. The locations of the centroid were computed from deposition-weighted means of the latitude and longitude of the interpolation points. Unlike the centroids, the orientations of the major axes were computed from principal component analysis using the screened observations at the ISDME sites. A discrepancy between the observed and predicted orientation would indicate that one pattern is rotated from the other.

The measure quantifying the discrepancies between the location of the observed and predicted maximum amounts (Element 2) considered the uncertainty in our ability to identify the actual location. This uncertainty arises from two factors: the unfavorable separation distances between the monitoring sites (approximately 200 km across the high deposition region of North America) and the uncertainties of the observations themselves. This measure was quantified by calculating the shortest distance between the location of the maximum interpolated prediction and the envelope containing all possible locations of the actual maximum amount, as determined from the uncertainty limits about the interpolated observations.

Likewise, the measure quantifying the discrepancies between the magnitude of the observed and predicted maximum amounts (Element 2), not necessarily at the same location, considered uncertainties. If the maximum interpolated prediction was within the uncertainty limits of the maximum interpolated observation, we could not state with confidence that the model failed to predict the magnitude of the maximum amount. For this case, the measure would have a value of zero. If it were not, the value of the measure would be the difference between the predicted value and the value of the lower or upper limit of the uncertainty band.

The last measure, which addressed Element 3, quantified the discrepancies between the mean temporal patterns of the predictions and observations. Temporal patterns were determined by averaging the relative seasonal contributions to the annual amounts of sulfur wet deposition at the 30 ISDME sites where both predictions and screened observations were available for all seasons. The value of this measure was determined for each season by computing the difference between the observed and predicted mean seasonal contributions.

Unlike the other elements, no meaningful measures were suitable for Element 4, namely, the reasonableness of the predictions of air concentrations and dry depositions. However, available data and intercomparisons of the predictions were used to identify peculiarities and major differences in model behavior.

MODEL PERFORMANCE RESULTS

When compared to the spatial patterns of seasonal and annual observations of sulfur wet deposition, the predicted patterns exhibited a much less detailed structure: the somewhat elliptical patterns were smoother and generally did not replicate the secondary maxima. In addition, the predicted patterns consistently anchored their seasonal and annual maximum amounts very near the high emissions region of the United States, across western Pennsylvania, Ohio and West Virginia. This contrasted with the migratory nature of the observed patterns, which seemed to reflect, to a much greater extent, the features of the precipitation patterns. The location of the observed seasonal maxima ranged from Ohio and West Virginia to southern Ontario. This contrast is illustrated by Figure 2, which compares the patterns of the interpolated summer observations and the MOE model predictions. In summary, the predicted patterns tended to reflect the emissions pattern, while the observed patterns showed a greater response to the precipitation patterns.

Measures of Performance

Both seasonal and annual patterns of observations and predictions were objectively compared by calculating the differences between the interpolated observations and predictions at each half-degree cell across eastern North America. Maps depicting those cells where differences were statistically significant were instrumental in distinguishing between regions of poor agreement and regions where predictions were within the uncertainty limits of the interpolated observations. Figure 3 provides an example of such a map. The percentages of the area within the entire evaluation region where differences were significant (column "ALL" in Table 3) indicated that across at least 80% of the evaluation region:

1) The seasonal predictions of five of the eleven models (ASTRAP, MOE, SERTAD, TVA and UMACID) were within the uncertainty limits of the observed patterns for every season;

2) For three seasons, the seasonal predictions of all but three models (RIVAD, RTM-II, and SIMPMOD) were within the same uncertainty limits; and

3) The annual predictions of all but three models (RIVAD, RTM-II and STATMOD) were within the uncertainty limits of the annual observed patterns.

A contributing factor to this high percentage was the degree of uncertainty of the observed pattern. The fact that the uncertainties were as high as $\pm$ 50% across the region of greatest site density, coupled with the fact that the observations varied in space by only a factor of two, the models were afforded a rather large statistical tolerance for differences.

As a whole, the predicted patterns replicated best the spring observed pattern, since for this season, the region of maximum observations most closely coincided with that of the predictions. Spring predictions of five models (AES, MOE, SERTAD, SIMPMOD and TVA) were within the uncertainties of the interpolated observations virtually everywhere.

Depending on the model, distances separating the predicted and observed locations of the maxima for each season were less than 170 to 350 km. Except for summer, most models were within the uncertainty limits of the magnitude of the seasonal and annual maxima. In summer, no model predicted the magnitude of the maximum within the uncertainty limits of the observed maximum; the maximum predictions for all but two models were less than the value of the lower uncertainty limit of the observed maximum.

The discrepancies between the observed and predicted centroids are a measure of the tendency of the models to spatially distribute sulfur mass differently than that which is observed. For each season, the centroids of most of the predicted patterns were within 100 km of those of the observed patterns; all were within 200 km. In general, the seasonal locations of the predicted centroids were to the south, southeast and east of the observed centroids, the locations of which were across or very near to northern West Virginia. The MOE model, which is the only ISDME model to simulate pollutant transport via geostrophic winds, consistently produced seasonal patterns with centroids to the north of that of the observed patterns. Therefore, unlike the other ISDME models, MOE exhibits a tendency to deposit more sulfur mass in the north, as opposed to the south, southeast and east, portions of the evaluation region.

Regarding the seasonal contributions to the annual observed amounts of sulfur wet deposition, nearly 40% and 30% of the annual observed amount occurred during summer (July, August and September) and spring (April, May and June), respectively. Approximately 15% occurred during winter (January, February and March) and autumn (October, November and December). As illustrated by Figure 4, most models were within 10% of the summer and spring contributions and within 25% of the autumn and winter contributions. Three models did not mimic the observed seasonal variability.

Two features of model performance were common to the four statistical models, but were not characteristic of many of the remaining models. First, the orientation of the major axes as determined by the statistical models, as well as ASTRAP and MOE, tended to be rotated more from the major axes of the observed pattern. That is, these models tended to produce patterns which were oriented more toward the east-west axis, as opposed to the southwest-northeast orientation characteristic of the observed pattern (Figure 5). In contrast, the orientation of the patterns produced by the deterministic models tended to be closer to that of the observed patterns. This distinction could be due to the fact that the statistical models base their transport simulations on seasonal mean wind vectors, the direction of which tends to be from the west. On the other hand, those models which use actual wind measurements (e.g., AES, RELMAP, RIVAD, RTM-II and UMACID) yield patterns with major axes more commensurate with those observed. The second feature common to the statistical models was the tendency to predict sulfate air concentrations higher than those of the other models at the ISDME sites as well as at the four Canadian sites where measurements were available.

Model Performance Scores

Finally, the seasonal, annual and overall performances of the models were quantified and intercompared via performance scores (Figure 6). These scores were based on the number of model performance measures passing the prescribed criteria, which are listed in Table 2. Since these criteria were based on the medians of the values of each measure, the performance scores should be interpreted only in a relative sense (i.e., for model intercomparisons). Overall model performance was based on a weighted mean of the seasonal performance scores. These weights reflected the relative seasonal contributions of the observed amounts of sulfur wet deposition to the annual total.

Figure 6 illustrates that three of the eleven models, MOE, SERTAD and TVA, performed as well as or better than the others for all four seasons; this is reflected in the overall scores for these models. As indicated by the overall scores in Table 4, these three models especially did well in predicting the amounts within the uncertainty bounds of the interpolated observed amounts.

A possible explanation as to why MOE performed better is that it was the only model to use hourly precipitation amounts; the other deterministic models used 12- or 24-hour amounts. Ley (1985) has concluded that by using hourly precipitation amounts, as opposed to 12-hour amounts, the MOE spatial patterns of sulfur wet deposition exhibited lower amounts across the higher emissions region and higher amounts downwind. This effect was also demonstrated by Coventry and Eder (1987) for RELMAP.

The SERTAD model appears to have performed relatively well for the wrong reasons. First, its background contribution to sulfur wet deposition is very high (e.g., across the high deposition region of eastern North America, 40 to 50% of the magnitude of its predictions was attributed to the background contribution). In addition, this model does not account for sulfate dry deposition and, consequently, predicts seasonal sulfate air concentrations that are greater than those measured by factors ranging from two to ten. Other reasons are discussed in the ISDME report (1987).

Two other models, ASTRAP and AES, had relatively high scores for summer and spring, but much lower scores in the two other seasons. For the autumn and winter, as indicated by Table 3, both models significantly overpredicted the sulfur wet deposition across large portions of the eastern United States. The seasonally constant scavenging rates of the AES model could explain its behavior. Meanwhile, the scavenging rates in ASTRAP could be too high, since this model tends to overpredict in every season.

The overall performance score for UMACID is similar to those for AES and ASTRAP. Unlike these latter two models, however, UMACID predictions significantly differ over a large area in summer and spring. In autumn and winter, UMACID predictions are within the uncertainties of the interpolated observations virtually everywhere. Similar to AES, the inconsistency in the seasonal performance could be explained by the seasonally constant scavenging rates used in UMACID.

The remaining five models, RELMAP, RIVAD, RTM-II, SIMPMOD and STATMOD, generally had lower seasonal performance scores, and therefore, lower overall performance scores. RELMAP significantly overpredicted in the highest deposition seasons (i.e., summer and spring). Further analyses have indicated that for summer and spring this model is very sensitive to the duration of the precipitation events. The three-hour durations used in the ISDME application resulted in significant overpredictions of the amounts of sulfur wet deposition across the region of high emissions. Based on analyses of hourly precipitation data, a more representative duration for these two seasons is one and one-half hours. When using this latter duration, the predictions are much more similar to the observations.

RIVAD and RTM-II, which are similar models, tended to significantly underpredict the interpolated observed amounts of sulfur wet deposition for every season. However, the models performed rather well for those measures not pertaining

to the magnitudes. Thus, there seems to be a systematic problem in the scavenging algorithms of these models. Also noted were much lower predictions of air concentrations of sulfur dioxide and sulfate and much higher predictions of sulfur dry deposition than those of the other models.

The seasonal performances of SIMPMOD and STATMOD, also a pair of similar models, were erratic. Although these models considered seasonally varying transformation, scavenging and dry deposition rates, little seasonal variation in the predictions of sulfur wet deposition was observed. As a consequence, these models tended to underpredict in summer and spring and overpredict in autumn and winter.

CONCLUSIONS

The new approach used in the ISDME to evaluate sulfur deposition models, which utilized performance measures focusing on seasonal pattern replication, has proven to be useful and is recommended for future evaluations of regional deposition models. This approach has determined for each season and the year i) the degree to which the observed patterns have been replicated and ii) those portions of subregions where interpolated predictions significantly differed from interpolated observations. The seasonal evaluations, compared to annual evaluations, provided a more rigorous assessment of model performance: Evaluations based on annual periods alone fail to test the ability of the models to mimic the strong seasonality of sulfur wet deposition. Furthermore, with the seasonal values of the performance measures, one is better able to understand the behavior of the models and consequently, better equipped to improve them.

The ISDME has indicated that the seasonal patterns of predicted sulfur wet deposition tend to reflect the pattern of sulfur emissions, while the seasonal patterns of observations showed a greater response to the patterns of precipitation. In contrast to the patterns of observations, the patterns of predictions were consistently anchored to the high emissions region of western Pennsylvania, Ohio and West Virginia. Although the magnitude of the maximum predicted amounts were generally within the uncertainty limits of the maximum observed amounts, the locations of the predicted maxima were always to the east, southeast or south of those of the observed maxima. Depending on the model, distances separating the locations of the predicted and observed maxima for each season were less than 170 to 350 km. Despite these differences, across at least 80% of the evaluation region, the seasonal predictions of five of the eleven models were within the uncertainty limits of the observed patterns for every season.

Based on the model performance scores, two statistical models (TVA and SERTAD) and one Lagrangian model (MOE) performed better than the other models. However, the SERTAD model appears to have performed better for the

wrong reasons. The scores also indicate that astatistical model can perform as well as a Lagrangian model.

Although the number of precipitation chemistry monitoring sites across eastern North America in 1980 was much greater than in previous years, the uncertainties of the interpolated observed amounts werestill rather high; these uncertainties, which are in part a function of site density, ranged from 30% across Ohio to approximately 100% across Quebec. By virtue of this level of uncertainty, and the fact that the seasonal amounts of sulfur wet deposition vary by a factor of two, the models were afforded a rather large tolerance for differences. This tolerance level will decrease for subsequent years when more sites were operating. The precipitation chemistry data from the additional sites plus i) air concentration data at many more sites than now exist and ii) empirical estimates of sulfur dry deposition amounts would increase the rigor of future model evaluations.

REFERENCES

Coventry, D.H. and B.K. Eder (1987) Private Communication.

Finkelstein, P.L. (1984) The spatial analysis of acid precipitation data, J. of Clim. and Appl. Meteor., 23: 52-62.

Fox, D.G. (1981) Judging air quality model performance - Review of the Woods Hole Workshop, Bull. Amer. Meteor. Soc., 62: 592-602.

ISDME (1987) The International Sulfur Deposition Model Evaluation, Final Report, In press.

Ley, B.E. (1985) Private Communication.

MOI (1982) Atmospheric Science and Analysis Work Group 2 of the U.S./Canadian Memorandum of Intent on Transboundary Air Pollution, Final Report.

UDDBC (1985) A unified wet deposition data base for eastern North America: Data screening, calculation procedures, and results for sulphates and nitrates, 1980, Final Report prepared by the Unified Deposition Data Base Committee for the Canadian Federal-Provincial Research and Monitoring Co-ordinating Committee, 79 p.

Voldner, E.C., A. Sirois, and T.L. Clark (1984) Data screening and calculation procedures for the North American precipitation chemistry data to be used in the International Sulfur Deposition Model Evaluation, Transactions of the Air Pollution Control Association/ American Society for Quality Control Specialty Conference on Quality Assurance in Air Pollution Measurements, Boulder, Colorado, October 14-18, 13 p.

Watson, C.R. and A.R. Olsen (1984) Acid Deposition System (ADS) for statistical reporting: System design and user's code manual, Final Report, EPA-600/8-84-023, U.S. Environmental Protection Agency, Research Triangle Park, NC.

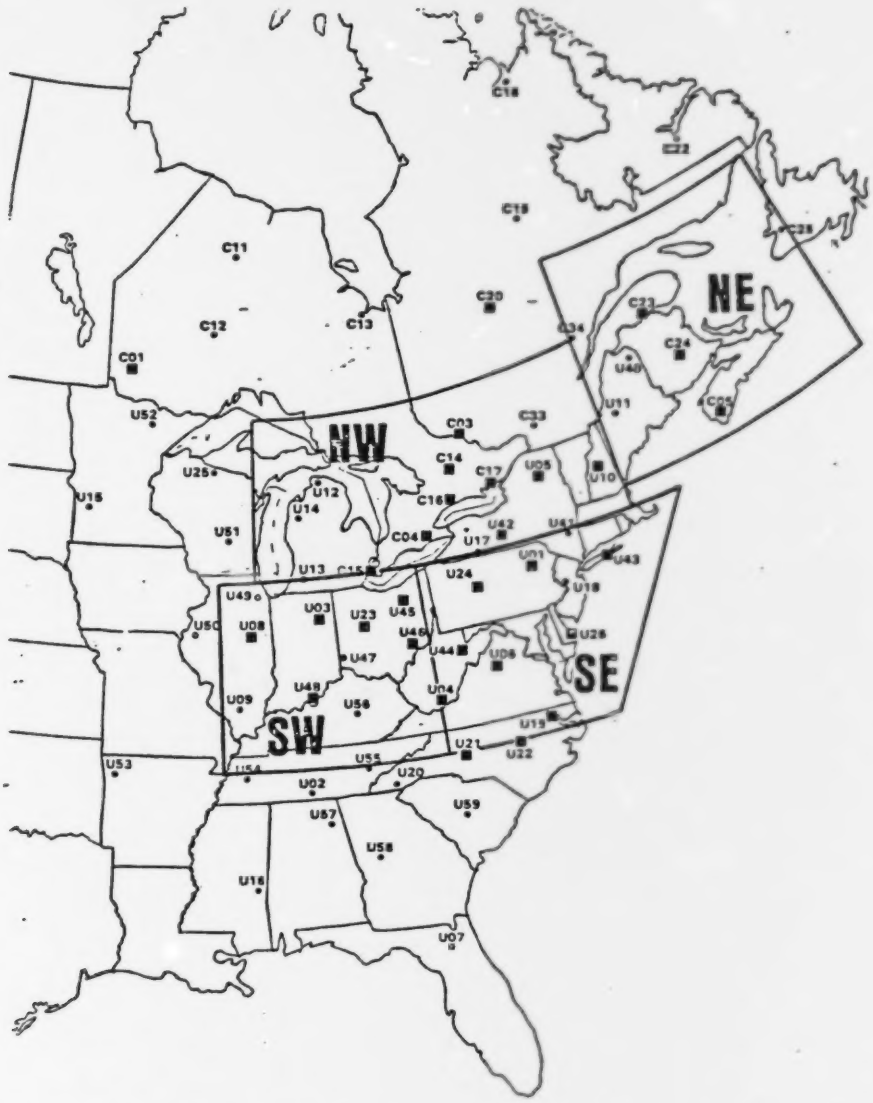

Figure 1. The ISDME sites and the four subregions of the evaluation region.

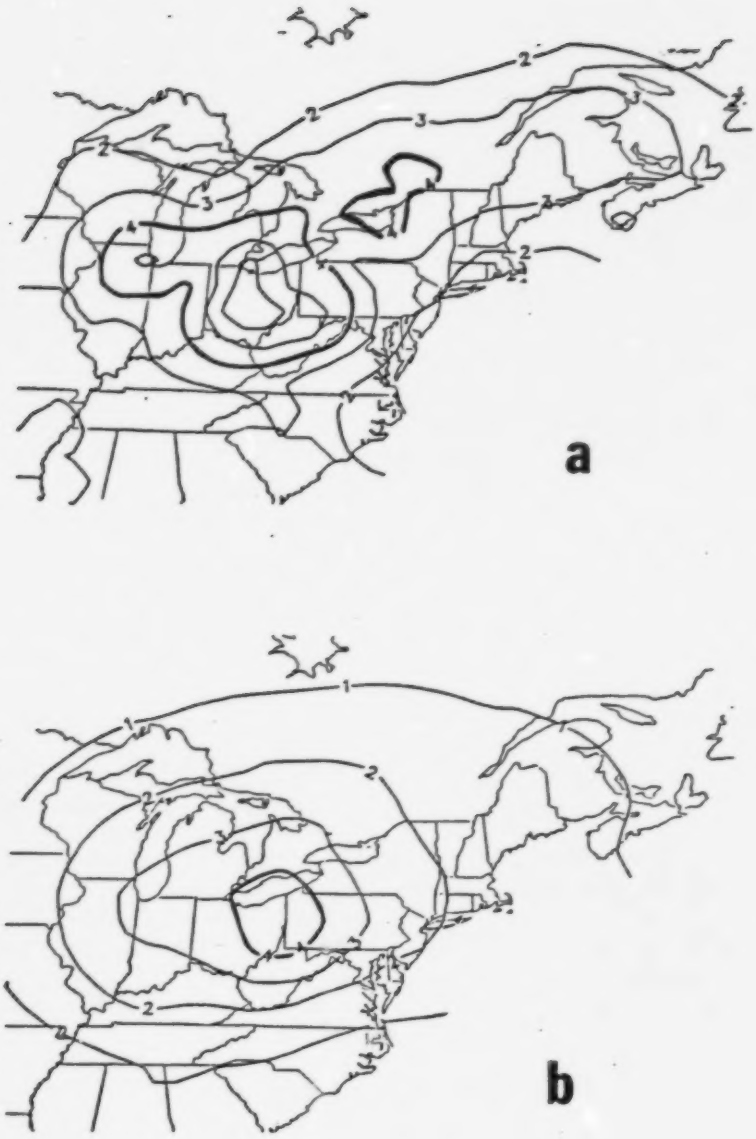

Figure 2. Spatial patterns of the interpolated observed (a) and MOE-predicted (b) amounts of sulfur wet deposition (kg S/ha) for the summer of 1980.

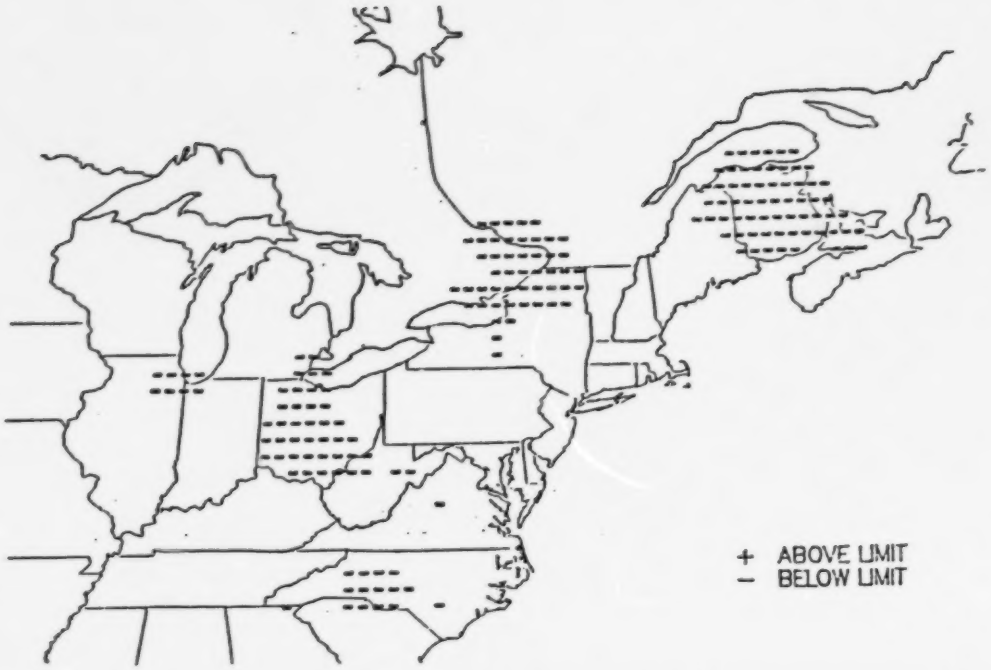

Figure 3. The half-degree cells where the interpolated MOE predictions significantly differed from the interpolated observed amounts of sulfur wet deposition for the summer of 1980.

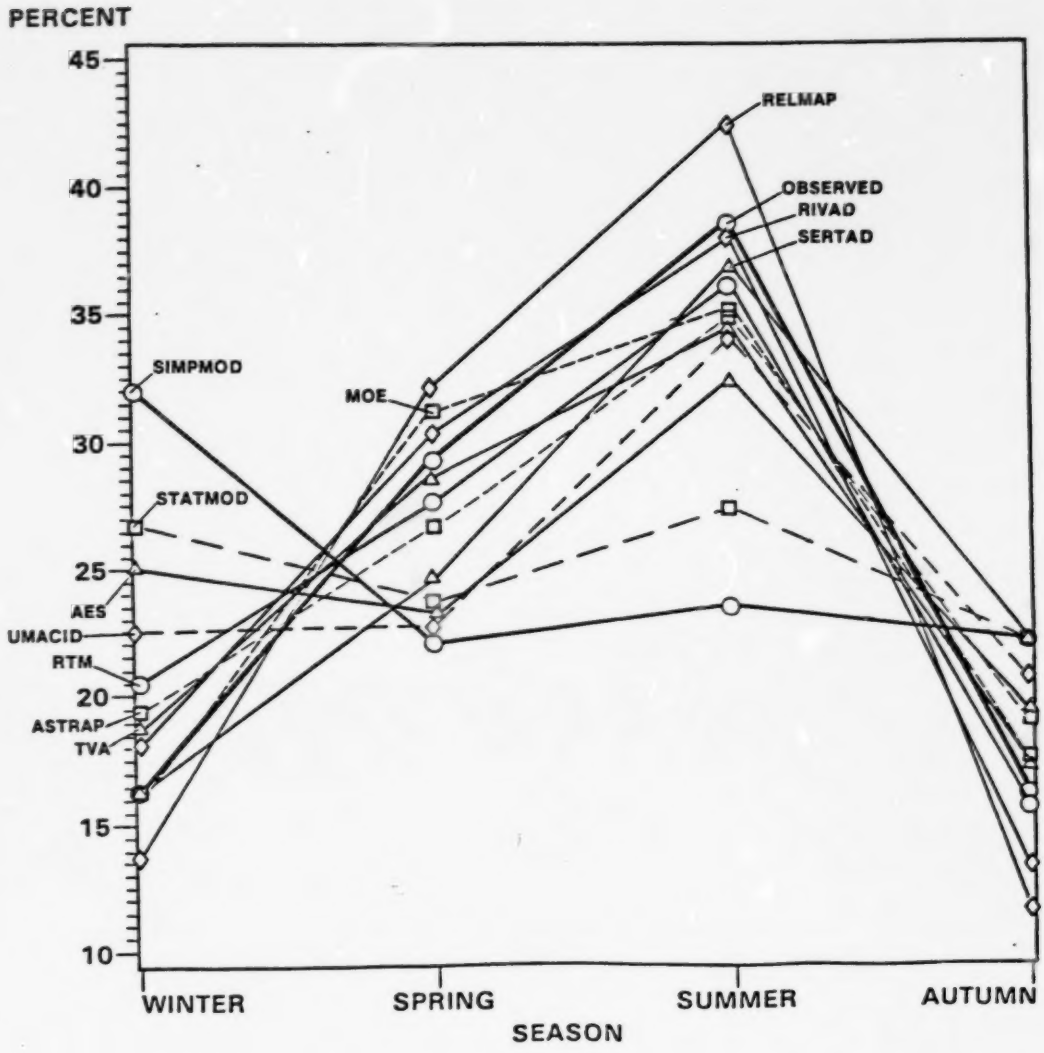

Figure 4. The 30-site mean of the relative seasonal contributions (%) to the annual amounts of sulfur wet deposition.

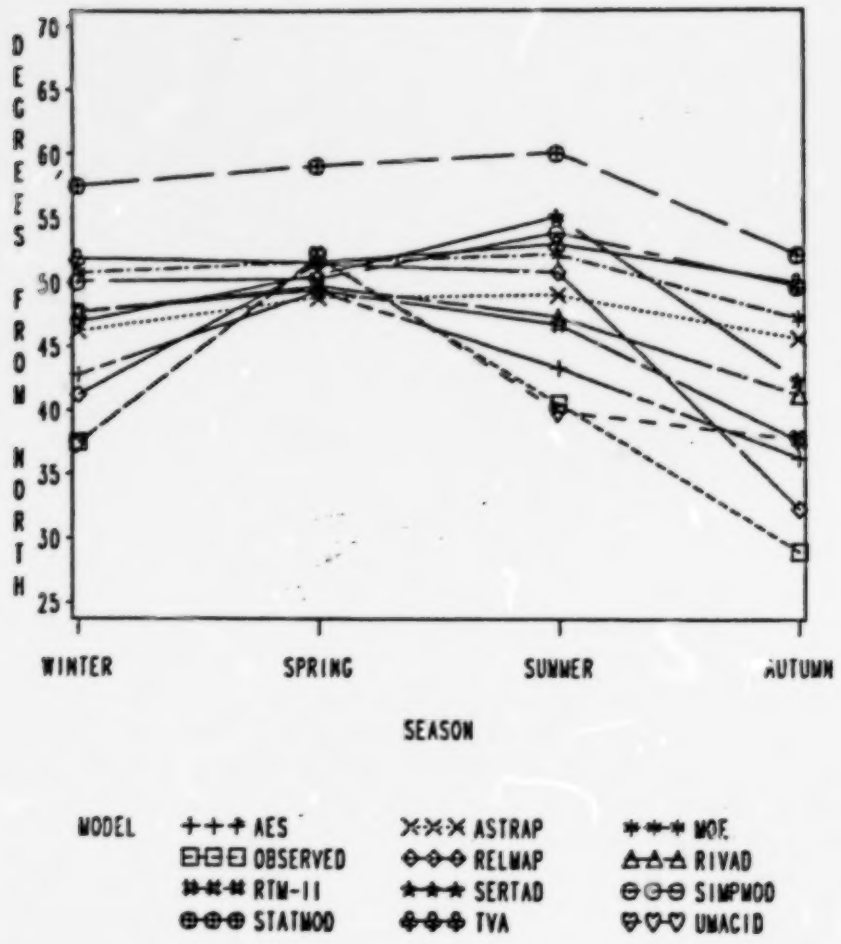

Figure 5. The seasonal orientations of the major axes of the site observations and predictions as determined from principal component analysis.

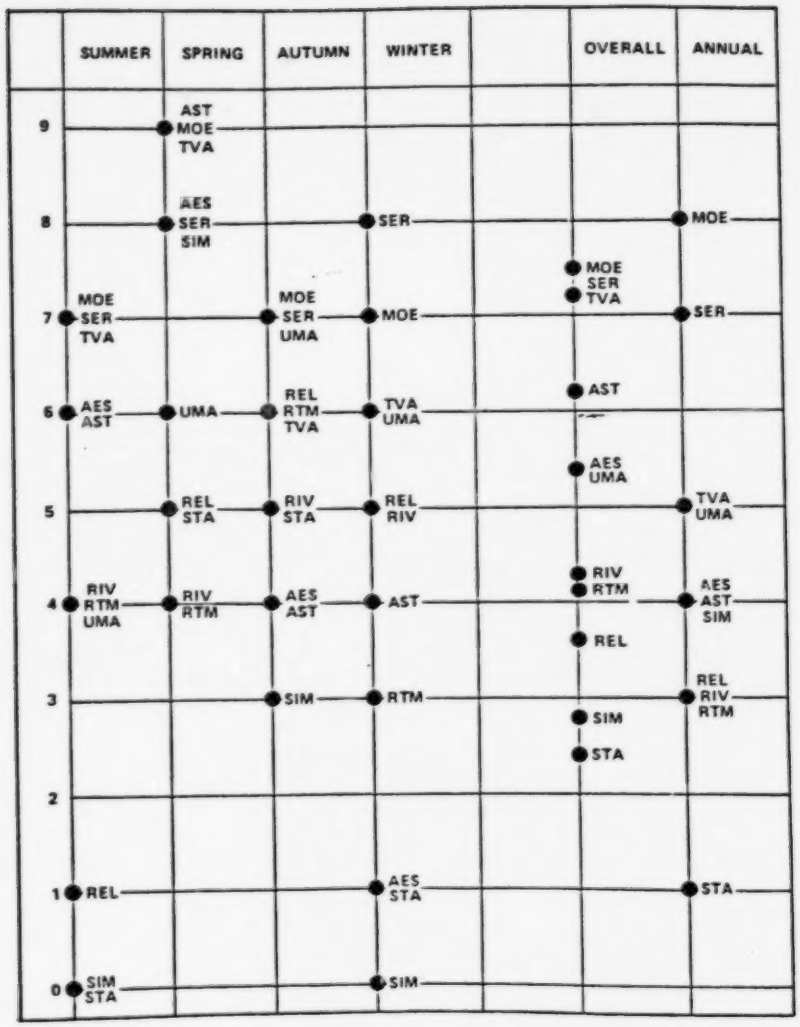

Figure 6. The performance scores of the ISDME models.

TABLE 1. THE ELEVEN MODELS EVALUATED IN THE ISDME.

ACRONYM	TYPE*	MODEL NAME
AES	L	Atmospheric Environment Service Long-Range Transport Model (Environment Canada)
ASTRAP	L	Advanced Statistical Trajectory Regional Air Pollution Model (Argonne National Laboratory)
MOE	L	Ontario Ministry of the Environment Long-Range Transport Model
RELMAP	L	Regional Lagrangian Model of Air Pollution (U.S. Environmental Protection Agency)
RIVAD	L	Regional Impacts on Visibility and Acid Deposition Model (Systems Applications, Inc.)
RTM-II	L/E	Regional Transport Model (Systems Applications, Inc.)
SERTAD	S	Model for Statistical Estimates of Regional Transport and Acidic Deposition (CONCORD Scientific Corporation)
STATMOD SIMPMOD	S	Statistical Model and another version (Enviornmental Research & Technology, Inc.)
TVA	S	Model T (Tennessee Valley Authority)
UMACID	L	University of Michigan Atmospheric Contributions to Interregional Deposition Model

* Model Types:

L	Lagrangian
S	Statistical
L/E	Lagrangian/Eulerian hybrid

TABLE 2. MODEL PERFORMANCE MEASURES, CRITERIA VALUES AND SCORES.

MEASURE	CRITERIA VALUE	SCORE
1. Area of Significant Differences	< 20% across evaluation region plus < 25% in each subregion	4
	OR	OR
	< 20% across evaluation region plus > 25% in one subregion	2
2. Centroids	< 75-km discrepancy	1
3. Major Axis Orientation	< 10-degree discrepancy	1
4. Location of Maximum	< 100-km discrepancy	1
5. Magnitude of Maximum	< 30% discrepancy	1
6. Temporal Pattern	< 15% discrepancy	1
Maximum Score		9

TABLE 3. PERCENT OF SUBREGIONS WHERE DIFFERENCES BETWEEN PREDICTED AND OBSERVED AMOUNTS OF SULFUR WET DEPOSITION ARE SIGNIFICANT. THE PREDOMINANCE OF UNDERPREDICTIONS (-) AND OVERPREDICTIONS (+) ARE NOTED.

	SUMMER 1980					SPRING 1980				
	NW	NE	SE	SW	ALL	NW	NE	SE	SW	ALL
AES	4-	17-	4+	15-	9-	1-	0	0	0	1-
ASTRAP	6+	1-	49+	5+	11+	2+	0	20+	0	4+
MOE	10-	18-	3-	17-	12-	0	0	0	0	0
RELMAP	6+	8-	90+	33+	24+	1+	0	60+	14+	12+
RIVAD	43-	26-	27-	47-	37-	16-	0	26-	58-	22-
RTM-II	48-	23-	26-	47-	39-	39-	1-	28-	64-	33-
SERTAD	3+	1-	4+	4-	3	0	0	0	0	0
SIMPMOD	39-	29-	3+	31-	30-	0	0	0	0	0
STATMOD	55-	38-	11-	30-	39-	33-	2-	0	2-	1%-
TVA	16-	22-	2+	11-	15-	0	0	0	0	0
UMACID	16-	27-	1+	31-	20-	13-	2-	3-	32-	13-

	AUTUMN 1980					WINTER 1980				
	NW	NE	SE	SW	ALL	NW	NE	SE	SW	ALL
AES	10+	2+	64+	0	13+	1+	0	86+	41+	21+
ASTRAP	15+	2+	54+	11+	16+	12+	0	53+	29+	18+
MOE	0	10-	0	0	2-	0	1-	5-	0	1-
RELMAP	4-	0	8+	0	3+	5-	0	51+	5+	10+
RIVAD	24-	21-	0	0	15-	31-	1-	0	4-	14-
RTM-II	19-	0	0	0	8-	28-	0	0	3-	12-
SERTAD	14+	11+	14+	0	10+	1+	0	20+	0	3+
SIMPMOD	3+	0	79+	18+	16+	15+	0	97+	62+	33+
STATMOD	1-	0	15+	0	2+	3-	0	33+	33+	12+
TVA	0	13-	11+	3+	5	0	0	13+	4+	3+
UMACID	0	0	3+	0	1+	1-	0	9+	2+	2+

TABLE 4. THE PERFORMANCE SCORES OF THE ISDME MODELS.

		Pattern Magnitude	Pattern Position	Maximum Value	Temporal Pattern	Total
SUMMER 1980	Maximum	4	2	2	1	9
	AES	4	1	1	0	6
	ASTRAP	2	2	1	1	6
	MOE	4	1	1	1	7
	REIMAP	0	0	0	1	1
	RIVAD	0	2	1	1	4
	RTM	0	2	1	1	4
	SERTAD*	4	1	1	1	7
	SIMMOD	0	0	0	0	0
	STATMOD	0	0	0	0	0
	TVA	4	0	2	1	7
	UNACID	0	2	1	1	4
SPRING 1980	Maximum	4	2	2	1	9
	AES	4	2	2	0	8
	ASTRAP	4	2	2	1	9
	MOE	4	2	2	1	9
	REIMAP	2	1	1	1	5
	RIVAD	0	2	1	1	4
	RTM	0	2	1	1	4
	SERTAD*	4	2	2	0	8
	SIMMOD	4	2	2	0	8
	STATMOD	2	1	2	0	5
	TVA	4	2	2	1	9
	UNACID	2	2	2	0	6
AUTUMN 1980	Maximum	4	2	2	1	9
	AES	2	1	1	0	4
	ASTRAP	2	1	1	0	4
	MOE	4	1	1	1	7
	REIMAP	4	1	1	0	6
	RIVAD	4	1	0	0	5
	RTM	4	1	0	1	6
	SERTAD*	4	1	2	0	7
	SIMMOD	2	0	1	0	3
	STATMOD	4	0	1	0	5
	TVA	4	0	1	1	6
	UNACID	4	2	1	0	7
WINTER 1980	Maximum	4	2	2	1	9
	AES	0	1	0	0	1
	ASTRAP	0	2	2	0	4
	MOE	4	0	2	1	7
	REIMAP	2	1	1	1	5
	RIVAD	2	1	1	1	5
	RTM	2	0	1	0	3
	SERTAD*	4	2	1	1	8
	SIMMOD	0	0	0	0	0
	STATMOD	0	0	1	0	1
	TVA	4	1	1	0	6
	UNACID	4	1	1	0	6
OVERALL 1980	Maximum	4.0	2.0	2.0	1.0	9.0
	AES	3.0	1.3	1.1	0.0	5.4
	ASTRAP	2.4	1.8	1.4	0.6	6.2
	MOE	4.0	1.1	1.4	1.0	7.5
	REIMAP	1.5	0.7	0.6	0.8	3.6
	RIVAD	1.0	1.7	0.8	0.8	4.3
	RTM	1.0	1.5	0.8	0.8	4.1
	SERTAD*	4.0	1.5	1.4	0.6	7.5
	SIMMOD	1.5	0.6	0.7	0.0	2.8
	STATMOD	1.2	0.3	0.9	0.0	2.4
	TVA	4.0	0.7	1.7	0.8	7.2
	UNACID	1.9	1.8	1.3	0.4	5.4
ANNUAL 1980	Maximum	4	2	2	-	8
	AES	2	1	1	-	4
	ASTRAP	2	1	1	-	4
	MOE	4	2	2	-	8
	REIMAP	2	1	0	-	3
	RIVAD	0	2	1	-	3
	RTM	0	2	1	-	3
	SERTAD*	4	2	1	-	7
	SIMMOD	2	1	1	-	4
	STATMOD	0	0	1	-	1
	TVA	2	1	2	-	5
	UNACID	2	2	1	-	5